

Internet Routing

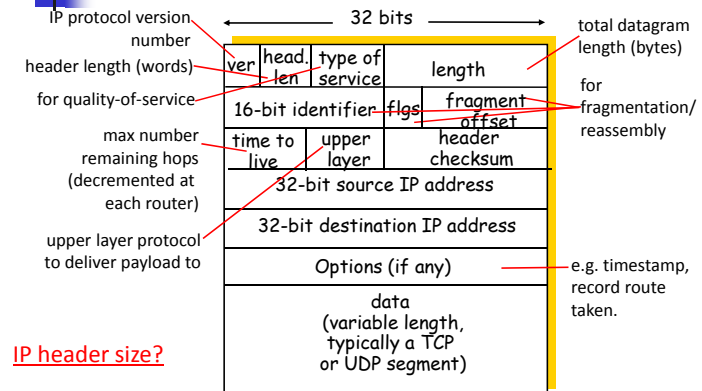
Kai Shen

9/29/2014

CSC 257/457 - Fall 2014

1

IP Packet Format



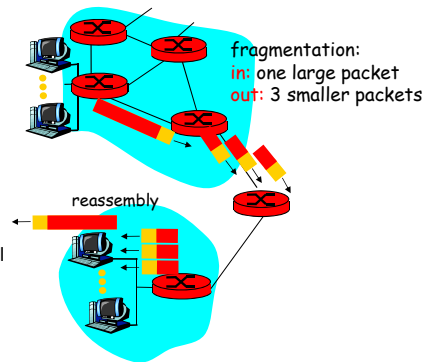
9/29/2014

CSC 257/457 - Fall 2014

2

IP Fragmentation & Reassembly

- Network links have MTU – largest possible frame.
 - different link types, different MTUs
- Large IP packet divided (“fragmented”) in network
 - one packet becomes several packets
 - “reassembled” only at final destination
 - IP header bits used to identify, order related fragments



9/29/2014

CSC 257/457 - Fall 2014

3

Principles of Internet Routing

Our routing study thus far - idealization

- uniform routing scheme across whole network
- uniform optimization goal
- ... does not work well in practice

scale: with hundreds of millions of destinations:

- can't store all dest's in routing tables!
- routing table exchange would swamp links!

administrative autonomy

- internet = network of networks
- each network admin may want to control routing in its own network

9/29/2014

CSC 257/457 - Fall 2014

4

Hierarchical Routing

- The Global Internet consists of **Autonomous Systems (AS)** interconnected with each other
 - Each AS is usually the network of one organization without internal conflict of interests
- Two-level routing:
 - Intra-AS**: administrator responsible for choice of routing algorithm within network, typically optimized for performance
 - Inter-AS**: unique standard to connect all ASes, performance may not be the only goal

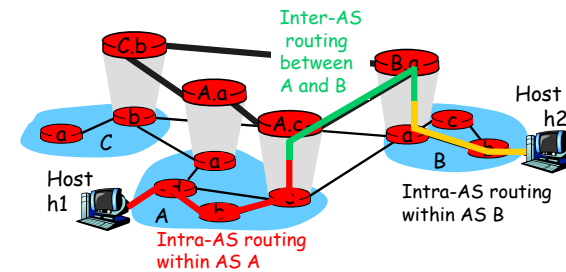
9/29/2014

CSC 257/457 - Fall 2014

5

Hierarchical Routing

One path can consist multiple route domains with different routing protocols

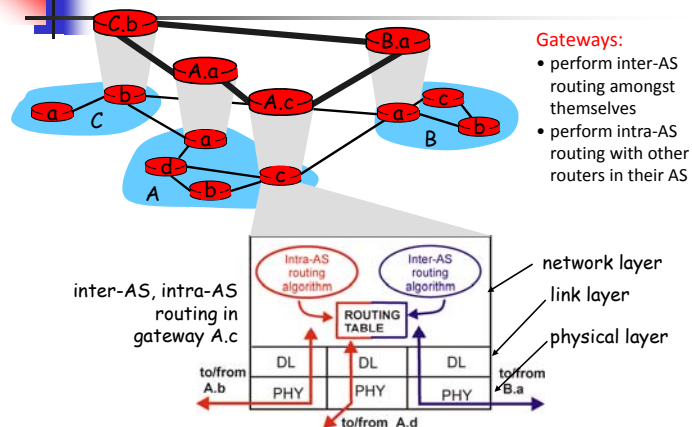


9/29/2014

CSC 257/457 - Fall 2014

6

Gateway Routers



7

Autonomous Systems in Practice

- Types of **Autonomous Systems (AS)**:
 - Stub AS**: small corporation: one connection to other AS's
 - Multihomed AS**: large corporation (no transit): multiple connections to other AS's
 - Transit AS**: provider, hooking many AS's together
- See a list of ASes

9/29/2014

CSC 257/457 - Fall 2014

8

RIP: An Intra-AS Routing Protocol

RIP (Routing Information Protocol)

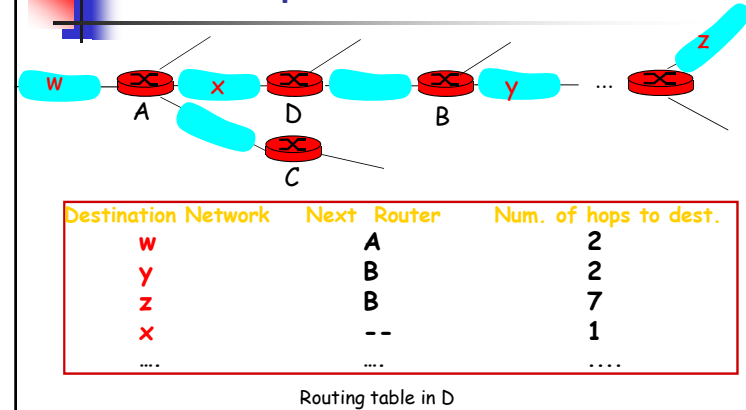
- Distance vector routing algorithm
- For scalability, each entry in routing table typically represents a network
- Distance metric: # of hops
- Routing tables exchanged (route advertisement) among neighbors periodically (e.g., every 30 secs)
 - why not stop after converging?

9/29/2014

CSC 257/457 - Fall 2014

9

RIP: Example



9/29/2014

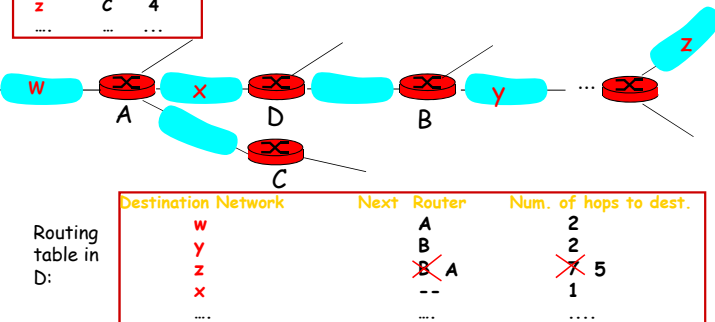
CSC 257/457 - Fall 2014

10

RIP: Example

Dest	Next	hops
w	-	-
x	-	-
z	C	4
...

Advertisement from A to D



9/29/2014

CSC 257/457 - Fall 2014

11

Routing Table Example

Check routing table at your own workstation:
`netstat -rn`

Destination	Gateway	Interface
128.151.67.0/24	0.0.0.0	eth0
169.254.0.0/16	0.0.0.0	eth0
0.0.0.0	128.151.67.100	eth0

- Local IP network contains two segments of IP addresses
- Routing tables in routers are more complex

9/29/2014

CSC 257/457 - Fall 2014

12

OSPF: Another Intra-AS Routing Protocol

OSPF (Open Shortest Path First)

- Uses Link State algorithm
 - Disseminate link state information to entire AS
 - Complete topology map at each node
 - Route computation using Dijkstra's algorithm
- Solve the scalability problem in OSPF
 - Hierarchical OSPF in large AS

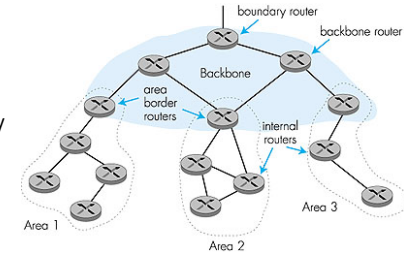
9/29/2014

CSC 257/457 - Fall 2014

13

Hierarchical OSPF

- Two-level hierarchy: local area, backbone.
 - Link-state broadcasts only in area
 - Each node has detailed area topology



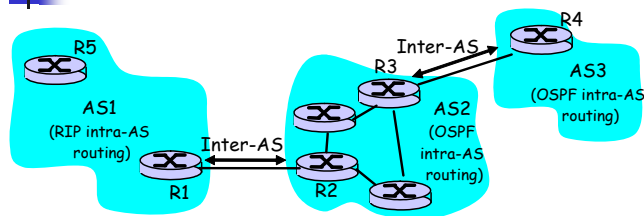
- Routers:
 - **Area border routers**: run intra-area OSPF routing as well as backbone OSPF routing.
 - **Backbone routers**: run OSPF routing limited to backbone.
 - **Boundary routers**: connect to other AS's.

9/29/2014

CSC 257/457 - Fall 2014

14

Inter-AS Routing in the Internet



- There should be just one inter-AS routing protocol
 - **BGP (Border Gateway Protocol)**
- Inter-AS routing
 - Across multiple administrative domains, which can have conflicting interests

9/29/2014

CSC 257/457 - Fall 2014

15

BGP Routing Protocol

- **Path Vector** protocol:
 - similar to Distance Vector protocol
 - each Border Gateway broadcast to neighbors (peers) entire path (i.e., sequence of AS's) to destination

Flexible routing management in BGP

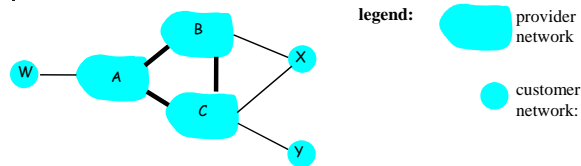
- A router selects routes offered by neighbors based on
 - cost
 - policy (don't route via competitors AS)
- A router can control incoming traffic by manipulating its route advertisements to neighbors:
 - e.g., don't want to route traffic from Z -> don't advertise any routes to Z

9/29/2014

CSC 257/457 - Fall 2014

16

BGP Routing Policy Management



- A,B,C are **provider networks**
- X,W,Y are **customers** (of provider networks), X is **dual-homed**
- X does not want to route transit traffic between B and C
 - ⇒ X advertises nothing to B or C
- A wants to help its customer (W) but not irrelevant traffic
 - ⇒ A tells W everything it knows
 - ⇒ A tells B,C about link to W but not link to C,B

9/29/2014

CSC 257/457 - Fall 2014

17

BGP Routing Table

- BGP routing table at an AS records complete paths to all other ASes
 - See an example
- Possible to construct (almost) complete internet topology (AS-level) with BGP routing tables from several routers

9/29/2014

CSC 257/457 - Fall 2014

18

Why two-level Intra- and Inter-AS routing?

Scale:

- Hierarchical routing saves table size, reduces update traffic

Policy vs. performance:

- Inter-AS: admin wants control over how its traffic routed, who routes through its network. Performance in a global sense is not a major concern for individual service provider.
- Intra-AS: single admin, can focus on performance.

9/29/2014

CSC 257/457 - Fall 2014

19

Disclaimer

- Parts of the lecture slides contain original work of James Kurose, Larry Peterson, and Keith Ross. The slides are intended for the sole purpose of instruction of computer networks at the University of Rochester. All copyrighted materials belong to their original owner(s).

9/29/2014

CSC 257/457 - Fall 2014

20